On the Spatial Allocation of Public Goods

Research Thesis

Presented in partial fulfillment of the requirements for graduation with research distinction in Economics in the undergraduate colleges of The Ohio State University

by

Robert Daniel

The Ohio State University

May 2019

Project Advisor: Professor Paul Healy, Department of Economics

1 Abstract

We wish to find an optimal solution to an economic problem concerning allocation of a public good by some central, governing body, over a spatial dimension. In particular, this government is cost minimizing, so it wishes to provide the minimal amount of the public good necessary to achieve its desired outcome, while at the same time considering its dual of maximizing the benefit of the limited amount of good which it provides. This paper puts forth a straightforward, easily-implementable algorithm using only some linear algebra and graph theory to solve this problem with suitable generality for varied applications in public economics and beyond.

In short, this algorithm, given a static, planar graph and a utility function solves for the constrained optimal placement of public good units so as to satisfy the above, using eigenvalue centrality and Fiedler partitions. These reduce large, difficult placement problems to those easily solvable using standard iterative methods.

We then validate this algorithm for fire hydrant placement in a real life neighborhood according to city codes and geographic properties. The algorithm performs accurately and does so at a polynomial time complexity.

2 Introduction

Spatial and network economics are both rather new phenomena in economic research, though network structures have long been studied in computer science, operations research, mathematics, and even sociology. Computer Scientists concern themselves with parallel computing, operations research focuses on planning and logistics, mathematics has a long history with graphs, and sociology cares about community detection and structure.

In its purest form networks are nothing more than applications of graph theory, founded by Leonard Euler in 1736 concerning the traversabilty of bridges in Konigsberg, specifically if it was possible to cross each of its seven bridges exactly once. This lead to questions concerning how does arrangement, placement, and general spatial and topological structures matter when formulating a problem mathematically. Since then, the study of graphs has evolved and lent itself to several areas of study with innumerable applications.

In general, a graph is a collection of vertices and edges, where the vertices represent the nodes and how these nodes are connected are encoded in the edges which link some two nodes together. Other structures are weightings on either nodes or edges, which rank some of these as more or less important, costly, different, etc.

than others. Some important questions in graph theory based on these simple notions are how connected is a node in general or in its neighborhood, do any cliques exist, how can we traverse along the graphs edges, how can we color the graph, how can we cover the graph, and how can we quantify the size of the graph.

The issue we wish to solve is a combinatorial optimization problem of finding where to allocate public goods in spatial networks given a utility structure with a punishment parameter based on distance. In short, the process is solving an extended vertex covering problem from graph theory, i.e. how to allocate goods such that all of the populace has strictly positive utility. The method works by finding the most central and connected node in a graph, checking the diameter for a cover, then if one is not found partitioning the graph until one is found, basically taking large graphs and chopping them up into easier to work with sub-graphs to guarantee a connected cover. The rest of this section serves to survey the economic literature concerning graphs, the literature review covers the methods contributing to the paper, and the model with an example then follows.

Spatial economics largely begins with the Tiebout model where agents are self selecting by sorting themselves into communities or groups such that their demand functions are met for local public goods. In a more concrete manner it can be realized as families selecting themselves into neighborhoods that provide child entertainment such as a pool whereas the elderly may sort themselves into communities which are in close proximity to a hospital which is what they have demand for. This is a foundational topic in the economics of space from which a vast literature has sprung forth. From this Henderson (1985) writes about the entrepreneurial implications of land use mechanisms in a single period long run equilibrium which can be extended to dynamical land use.

In the semi-recent economics literature, Jackson and Wolinsky (1996) formulate the standard, game-theoretic notion of pairwise stability in graphs and graph evolution. This model allows for the possibility of an equilibrium in graphs to be reached, that is how do agents, represented by vertices in a graph, create and sever links with another in order to gain the highest utility and Watts (1999) writes about how this dynamic process converges and if it converges in the first place, i.e. does the graph cycle through finite states without end. To do this they construct sequences of graphs called improving paths where each player may move once per turn to benefit himself based on his current graph structure. However, much of the strategy is removed as it is assumed players are not forward looking and maximize their utility solely on what exists at the present time.

A very simple example is: given two nodes with no edge between them and a utility structure in place, they will create a connection between them if and only if they both gain positive utility from creating this link. From this paper came Jackson and Watts' (2002) paper about a stochastic and strategic version of the above papers, concerning probability distributions of the degree of nodes to see how sequences of random graphs behave, cycle, and converge with applications to marriage matching and college admissions. Additionally, they define simultaneous improving paths which allow for two actions to be taken by each player each turn such as both deleting an edge and adding one, which they then extend to a stochastic environment to show the strength of their results and this stability notion.

Additionally, networks also play a large role in industrial organization. Shakkottai and Srikant (2006) write about competition among internet providers modeled using graphs to show variations in price with location. They find that these businesses merge at the local level to form some level of oligopoly or monopoly for which some level of collusion would be enacted so that prices and service became very similar if not identical. Moreover, transit ISP's, those that switch customers from different networks to optimize traffic would often be vertically integrated into the local ISP's so as to reap more or all of the profit.

Kranton and Minehart (2001) construct a graph theoretic formulation of buying and selling, and how those agents connect with one another and its economic outcomes. More specifically, they show it is optimal for these networks to form in the presence of uncertainty as buyers may not know the quality of goods provided or the consistency of the sellers and thus have a network demand to prevent being cheated or overly reliant on one seller.

In the realm of public economics, Bramoullé and Kranton (2006) study network structure and equilibrium of non-excludable public goods. In particular, they find encouragement of specialization in public goods as the pairwise stable equilibrium often encourages free riding on a neighbors specific contribution, e.g. when someone plants a garden, all their neighbors do not also plant one and simply reap the benefit of that person's garden which then allows another neighbor to instead build a pool which others share in and so on. Also they show that there exist large positive externalities from these goods as benefit permeates throughout the network, but welfare decreases when networks become too dense as free riding becomes too attractive and the benefit gained from providing a good becomes too low.

Labor economics also has rich uses of graph theory by relating employment processes to graph based social networks. Davern (1997) writes about informal job searches where personal recommendations from social networks of sufficient size allow for easy job attainment. More thoroughly, Montgomery (1994) writes about Markov chains in employment matching using a notion of weak and strong ties where agents use both inter and intragroup social actions to maximize the chance of employment given that they are not competing for the same jobs as those in their own groups.

In social networks, Tichy, Tushman, and Fombrun (1979) in a survey about studying the inner operations of organizations, provide an overview of measuring influences, reputation, decision, and interactions using statistical methods. Newman (2001) provides an important paper on author collaboration using graph theory concerning how different scientists in different fields collaborate and what it says about the structure of how research is done and academic success.

Similarly, in mechanism design, is a seminal paper by Groves and Ledyard (1977) which attempts to solve public good allocation with behavior tempered by a punishment parameter which must be optimized to produce efficient levels of the public good in a decentralized manner. In general, Besley and Coate (2003) approach the problem from the side of political economy studying if decentralized mechanisms are a reasonable solution or if legislative processes are necessary considering the overlapping and external effects of any policy or mechanism.

In a more explicit spatial context, Szolnoki and Perc (2016) test mechanisms in spatial public good allocation and behavior of economic agents and how they impact the equilibrium with "tolerance" and adverse interactions between groups of cooperators, defectors, and loners. They examine the global stability of such games and the survivability of local solutions implying global solutions are in fact conglomerations of many smaller scale solutions making the "true" solution "hidden". Additionally, Wang and Chen (2018) solve mechanisms stemming from Groves and Ledyard for second order free riding, that is free riding on others enacting punishments, with spatial public goods, finding an optimal level of peer punishment that is a dominant strategy implementable given sufficient population size and mixing.

Econometrically, Sieg et al (2004) estimate total benefit of spatial public goods in environmental markets for air quality and pollution. They find clear policy implications stemming from the fact that housing markets are indeed impacted by changes in air quality under different levels of mobility. Computationally, Athreya and Somanathan (2008) estimate create an algorithm to assess the optimality of India's post office placement. They approximate losses due to India not minimizing the median distance between post offices and thus causing welfare losses especially among the under-served and poorer populations. Talen (1998) assess the placement of local public playgrounds with respect to public transportation and family traversabilty considering an urban planning approach to the problem of providing adequate resources for children of impoverished families.

In the geographic information science literature, Cha (2008) writes about optimal public transportation using origin-destination models with proximity-cost measures to see if those who are disadvantaged have adequate access to city features and its policy implications in places such as labor markets and general accessibility. In another study concerning just labor economics, Hsieh and Moretti (2015) study land use regulations in the restriction of housing supply and its effects in dynamic labor markets which cause spillover into adjacent communities due to insufficient housing and high demand for employees.

3 Literature Review

In the more closely related literature, We will survey a few topics of particular importance for this paper's algorithm formulation and papers dealing explicitly with solving spatial economic systems with emphasis being given to those with algorithmic or computational components.

Newman (2004) survey algorithmic network analysis in a seminal paper motivating the research done in this paper. They review spectral partitioning such as Fiedler's partition and the Kernighan–Lin algorithm, sociological hierarchical clustering and other topological data analysis, greedy algorithms for edge removal, modularity, and physical electric network methods reformatted for different kinds of data. One particularly interesting approach is the last by Wu and Huberman (2004) which is both fast and accurate, if not unorthodox.

Wu and Huberman (2004) use Kirchhoff's loop rule and voltage vectors, due to the applications of their non-limiting assumptions, to first solve implicitly for the graph Laplacian using source and sink detection and checks how each node linearly combines with others. The original source and sink are the two nodes farthest apart, geodesically, as they are generally the least likely to belong to the same community and then builds off of this iterating through the other nodes to find likely communities. When applied to the standard karate club network it is able to correctly detect the community structure underlying it as validation. Raghavan, Albert, and Kumara (2007) improve on the above algorithm by lessening the amount of assumptions required in that the number of communities existing need not be known using unsupervised learning.

Algebraic Connectivity of Graphs by Fiedler (1973) is one of the most important papers for informing this paper and much of the theoretical computer science literature. It proves large amounts of assertions for simple connected graphs of which the most important is about the second smallest eigenvalue of a matrix M which is symmetric, positive semi definite, is strictly greater than zero and is bounded by the vertex degree set. Moreover it governs the edge deletion set necessary to bisect a graph into two non overlapping subgraphs which are themselves connected. The importance of this is explained more clearly by Slininger (2013) who likens the second smallest eigenvalue and its associated eigenvector as a period on a vibrating string with one half below the x axis and one half above just as the way the eigenvector is split and the partition is done-along the sign of the elements of the eigenvector. In particular, Fiedler's spectral partitioning uses the second smallest eigenvalue of the graph Laplacian, a matrix representation of the edges and vertices of a graph, to govern how the graph should be split with the elements in the eigenvector above zero going to one sub-graph and the negative elements going to the other.

Other algorithms applying the Fiedler Partition are numerous and surveyed as follows: Grady and

Schwartz (2006) apply this to the image segmentation problem who use graph methods on pictures to find their main components and object boundaries in linear time. Barnard and Simon (1994) and Driessche and Roose (1995) define an algorithm for approximating the Fiedler vector rather than calculating it explicitly using graph contractions where an edge is deleted and the nodes it connected to are made into one vertex, allowing for simpler calculations while still keeping most of the graph's structure. Chen and Hero (2014) uses Fiedler Partitions with topology and machine learning to detect real world community structures using "big data" from social media.

Additionally, Qiu and Hancock (2005) contribute heavily to the formulation of this paper, providing a method using Fiedler Partitions to decompose difficult graph problems into smaller sub-problems. They use this to cut a graph into non-overlapping sections and match sections using probabilistic methods to determine neighborhood structure of sub graphs. Then they apply this to a hierarchical clustering which analyses the metric and topological properties of the graph and sub-graphs. They prove this algorithm is numerically stable and accurate, and then apply their algorithm to imaging software which allows them to more accurately assess corners of toy houses and then assess structural similarity between different kinds for algorithmic validation.

Another major method of this paper is eigenvector centrality measured in a way similar to the process Google's Page Rank used at one point. Defined more thoroughly in the model section, Newman (2006) writes about how to more concretely and computationally to treat the problem of centrality and connectedness in the increasingly complex graphs and networks in places like google's search engine which has to iterate through possibly billions of different web pages. The underlying process uses linear algebra to solve for a centrality vector based on the underlying graph structure given by the adjacency matrix, a 0-1 matrix representing the edges of a given graph.

Other centrality measures are posed by Freeman (1977) who uses betweenness of nodes as a measure of centrality in sociological communication networks. He defines betweenness as being a node that most often falls in the middle of a group of nodes shortest paths/geodesics, and thus it must be central as it is required for optimal traversabilty. From this a second paper by Freeman (1978) takes a look at a complete measure of sociological centrality and considers degree centrality, that is the number of edges a vertex emits and closeness centrality by the weighted sums of adjacent nodes.

On assessing the diameter of a graph, that is the longest of the shortest paths between any two nodes, Seidel (1992) solves for an efficient implementation of the all pairs shortest path problem. The algorithm solves for a distance matrix given only the adjacency matrix of a graph given that it is undirected and unweighted. For their weighted counterparts, Aingworth et al (1999) devise an efficient algorithm which approximates the diameter to avoid costly matrix methods and with high accuracy. In general for shortest paths, Dijkstra's (1959) algorithm is the seminal paper spawning innumerable papers improving on his formulations. The basic formulation is it checks edge weight distances through nodes, flags the node as checked and then iterates through the others until all nodes are checked, from there it calculates the minimum path. Jonker and Volgenant (1986) modify Dijkstra's algorithm to use linear assignment which has initialization and augmentation phases for this linear program, with the first row reducing and the second permuting rows and columns iteratively to find the shortest path.

This algorithm, presented in the next section, effectively glues all of the ideas presented here together. It partitions large difficult graphs into tractable sub-graphs, finds the centrally located vertex, and then checks the sub-graph's diameter to determine if a covering has occurred. This algorithm solves a spatial allocation problem, well known in economics, using well researched methods from mathematics,operations research, and computer science.

4 Model and Theory

Our problem, which we wish to address, is can we design a system which efficiently allocates public goods with a spatial dimension. This can be thought of as where to place a public park(s) so as to maximize societal benefit but minimize cost. In particular, we wish to have this public good "cover" that is it helps, to some degree, everyone around it, such that no one has some non-positive benefit. Additionally, we wish to constrain it such that the placement(s) gives benefit maximally, that is no other spot will generate as much benefit as the location(s) which we put it.

First we wish to quantify how people are related spatially. To do this we must consider a simple, connected, finite Graph, G(V,E), where V is the vertices and E is the edges. From there we define the graph using the adjacency matrix, A, which is defined as follows.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$a_{ij} = \begin{cases} 1 & \text{if an edge exists between i and j} \\ 0 & \text{otherwise} \end{cases}$$

Thus we have a zero-one matrix that is strictly zero on the diagonal, as we assume that no node is connected to itself. Now we define the utility structure which we wish to assess, for any agent given by:

$$\forall i \in I : u(i) = b - g(j)^p \geq 0$$

That is, we allow the public good to exert non negative benefit for person i, given the public good is located at person j, with g being the geodesic distance (minimum edge traversal needed) between i and j. Then $p > 1$ is to indicate that utility decays rapidly as distance increases. We restrict ourselves to positive utilities only as we assume this public good provides only benefit and being located far away from the good does not affect one's utility.

Now we must decide where to begin looking for optimal placement of this public good. A natural thought is to place it with the person/node which is best connected. To be best connected we say it has some measure of having many adjacent vertices and those vertices themselves being adjacent to many nodes too. Thus we consider the problem, popularized by Google's page-rank posed most succinctly by Newman (2006) as:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} a_{ij} * x_j$$

which when converted to matrix and vector form is clearly the eigenvalue problem

$$Ax = \lambda x$$

$$det(A - \lambda I) = 0$$

where any λ is a solution, x is our centrality vector, and lambda is our vertex weighting.

Here we consider the largest eigenvalue which by the Perron-Frobenius Theorem is guaranteed to have an associated eigenvector that is positive and real. Then we can compare elements in the eigenvector corresponding to nodes on our graph G to one another without taking the norm or comparing real and complex numbers.

Thus the most connected node is the element in the eigenvector which has the largest value and is a reasonable node to choose for placement. However, while this does optimally benefit the graph in that it allows utility to reach the maximum number of nodes, it does not guarantee a covering such that no node is left with zero utility, for most graphs of reasonable size. While we could iterate through our eigenvector element listing this brings us to another problem of cost minimization. The government only wants to build the minimum amount of the public good to cover and no further. The point which is maximally connected and central is often located very close to other connected and central points meaning that those which are peripheral to our first choice will similarly be peripheral to our second choice and so on.

Moreover for a more general setting we wish to define a heuristic for which to reasonably solve large graph problems of this kind. Therefore we wish to impose a sort of partition on the graph so that we can work with smaller, more easily solvable graphs. That is we wish

$$minimize \sum v \in V$$

the sum of the nodes chosen, for which it is also a covering

A solution to this problem requires the Laplacian Matrix, L, which is defined as follows:

$$\mathbf{L} = \begin{pmatrix} d_{11} & -a_{12} & \cdots \\ -a_{21} & d_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$l_{ij} = \begin{cases} -1 & \text{if an edge exists between i and j} \\ 0 & \text{if there is no edge between i and j} \\ d & \text{if i=j} \end{cases}$$

where d is the degree of node i, that is the number of edges it has.

Then we consider the Fiedler Partition, that is an equitable bisection of the graph such that the maximum number of edges are preserved. This relies on the spectrum of the Laplacian which is formulated with Lagrange multipliers by Holzrich et al (1999) as: minimize f(x), given that $g_1(x)$ and $g_2(x) = 0$

$$\nabla f(x) - \lambda_1 \nabla g_1(x) - \lambda_2 \nabla g_2(x) = 0$$

where

$$f(x) = \frac{1}{2} x^T L x$$

$$g_1(x) = e^T x$$

where e^T is the unit vector

$$g_2(x) = \frac{1}{2}(x^T x - n)$$

where n is the number of nodes. This then can be reduced to the eigenvalue problem:

$$Lx = \lambda_2 x$$

$$det(L - \lambda_2 I) = 0$$

As this partition preserves connections it keeps connected "cliques" together and removes that which is outside the clique or otherwise peripheral. To do this we follow the writings of Fiedler and consider the second smallest eigenvalue that solves the above equation and its associated eigenvector. This eigenvector has elements that are then divided along their sign with those being greater than zero belonging to one sub-graph, G1, and those being less than zero belong to the other, G2.

Next we consider a "pseudo-greedy" algorithmic plan to optimize each sub-graph in an attempt to optimize the whole graph. As it is well know that a classic vertex covering is NP hard by Karp and Cooke, we wish to extend this vertex covering notion to look if benefit permeates throughout the graph, e.g. if node i is two edges away from node j, who has the public good, does node i have positive utility? For large benefits, small p weights on our geodesic distance, or a combination of both, the problem becomes simpler as it eliminates many possible choices needing consideration.

Under the assumption of homogeneity of benefit structures and permeability, we find the maximal reach possibly attained by the simple discrete maximization of our utility formula fixing p, i, and j then solving for the geodesic.

This gives us our constrained graph diameter, defined as the maximal number of edges needing traversed using the shortest path from node i to node j, where node j is fixed.

$$d = max(g(i, j)), \forall i \in I$$

We then combine this diameter notion with the λ_{pf}'s largest eigenvector element being our fixed j. It is then easy to check if the diameter of G with fixed node j exceeds our tolerance using standard numerical methods such as Dijkstra's Algorithm.

Then if the diameter exceeds our tolerance we partition the graph using Fiedler's Method. Then for the two sister sub-graphs it produces, we consider the adjacency matrix of both and re-solve for each of their

λ_{pf}'s, and check the diameter of the sub-graph with respect to the central node on the sub-graph.

If either sub-graph again exceeds d then we check if combined with its sister sub-graph's λ_{pf} in our partition, it creates a maximal covering, that is can this sub-graph be maximal with respect to its other half of the partition. If this also fails then we iterate this, partition the sub-graph(s) until we reach a diameter of sufficient size, and then check it with its sister.

This inter-step checking serves an important purpose in ensuring a minimally overlapping vertex covering. Fiedler's Algorithm only works on the discrete level and as such cannot always guarantee a perfectly equitable partition. As such, after several iterations it may produce a mixed geodesic sub-graph, i.e. part of the sub-graph has geodesic distance of at most d but there exists at least one vertex that exceeds this d. Thus it is necessary to check if this partition, when accompanied with its sister sub-graph, has a covering and thus a stopping point or if it requires an additional partition. An equally valid and less computationally expensive approach would also be to simply check the final answer iteratively to see if any answer is not necessary for producing a covering. That is, we could delete one answer and check for a covering and flag if there is still a covering and then proceed to the next solution. As shortest path algorithms are well developed this will not present a problem.

Together, as the adjacency matrix guarantees the best connected and most central node, Fiedler's partition guarantees the best partition possible, and inter-step/final-step checking eliminates local optimal, non globally optimal constructions, we can guarantee that this is a reasonably efficient "pseudo-greedy" algorithm. More specifically we have:

$$max \sum_{1}^{n} u(i),\ min \sum v \in V$$

chosen, such that

$$\nexists i \in I, u(i) = 0$$

In short, for large graphs with many possible vertex combinations needing to be considered, a brute-force method quickly becomes intractable. This algorithm splits our problem into many small and easily solvable sub-problems and allows for flexibility in permeability of benefit flow. However its correctness is bounded by the correctness of the Fiedler Partition which then serves as an upper bound to the accuracy of this problem making it an approximation algorithm. Currently the only way to guarantee perfect accuracy is to check all combinations which is of the order n^n which becomes impossible to solve on any computer for sufficiently large n.

Now it is necessary to consider the time complexity of the algorithm and its feasibility in implementation. No computation in this algorithm ever needs to be exact as magnitude and direction are the only necessary components. Similar to google's Page Rank algorithm we also use the power iteration for finding the largest eigenvalue and its corresponding eigenvector at time complexity n^2. Next it is necessary to consider the eigenvalue and vector for the Fiedler partition which Locally Optimal Block Preconditioned Conjugate Gradient does well with complexity n, as we only need one eigenvalue and vector which corresponds to its smallest non zero eigenvalue making it more easily solvable. Finally, computing the eccentricity/diameter of the graph is of time complexity n^2 by Dijkstra's algorithm. All together this gives us average time complexity of $O(n^4 + n)$, which while sub-optimal is not a means for concern due to the bounds on the size of the graph given its best use in at most the local government level. Additionally, as explicit calculation of most of the Fiedler process is superfluous and as such the complexity can almost certainly be optimized to be reduced to $O(n^3)$ or less.

4.1 Non-Homogeneous Utility

While the above section deals solely with homogeneous utility this constraint is not necessary and the model can be easily extended to a large distribution of graph utilities. In particular, our calculations can be adjusted so that those who benefit less from the public good and those who benefit more from the good given any geodesic distance can be accounted for, defined as

$$\forall i \in I, u'(i) = b(i) - w * g(j)^p \geq 0$$

where w is our individual weighting.

In modern semi-definite programming, the Fiedler partition can be extended to any non negative, real, symmetric matrix. However we wish to store the information about node weights separately from our usual matrices, in particular we wish to store w on the nodes themselves. The reason for this is it allows to see how the decay parameter is changed, e.g. for w equal to one half a larger geodesic distance is allowed for the node as its utility goes to zero less quickly. On the other hand, a weighting of w at 2 makes the geodesic distance twice as punishing and thus requires the node to be closer to the good to reap any utility.

Thus we define our node weight vector as the following:

$$n = [n_1 n_2 ... n_n]$$

where each n_j is the person's individual weighting. Then we solve for each person's maximum allowable distance, that is their distance from the public good where there utility is strictly positive.

$$max : g(j)$$

such that

$$b(i) - w * g(j)^p > 0$$

which we then store in an allowable distance vector in the same way as the node weights, and is easily solvable as b(i), w, and p are already known.

Next we proceed similarly to the above section where a central node is found, the graph's diameter is found and if it is larger than the maximum allowable distance of any node the graph is partitioned using Fiedler's method. The process than repeats until a suitable diameter is found such that the allowable distance vector is satisfied for the given sub-graph.

4.2 Edge Weights and Uncertainty

Next we introduce a notion of non uniformity of connections that is some edges can be weighted to be more or less "valuable" than others. These take two main forms either deterministic or stochastic. The former could indicate some measure of cost modification such as difficulty in traversing a particular route, in terms of the above example it could represent that a family struggles to go uphill to get to the public pool at a certain node across a certain edge. Stochastic weighting indicates uncertainty if an edge exists in the first place. An example of this could be seen in job retraining networks where a government wants to evaluate, given limited resources, which of the unemployed people could be retrained to provide maximal benefit to a community measured through meeting the demand of the people in this network.

This, which is similar to node weighting, has a simple modification

$$\mathbf{A}' = \begin{pmatrix} a'_{11} & a'_{12} & \cdots \\ a'_{21} & a'_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$a'_{ij} = \begin{cases} weighting \in R^+ & \text{if an edge exists between i and j} \\ 0 & \text{otherwise} \end{cases}$$

Then we can define the modified Laplacian as follows

$$\mathbf{L}' = \begin{pmatrix} l'_{11} & l'_{12} & \dots \\ l'_{21} & l'_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$l'_{ij} = \begin{cases} -weighting & \text{if an edge exists between i and j} \\ \sum weightings & \text{if i=j} \\ 0 & \text{otherwise} \end{cases}$$

Similarly to the above, since all weights are positive, both the Fiedler method and centrality methods perform the same as nothing changes the semi-definite nature of the matrices.

5 Validating Example

We validate the algorithm using real world data from a local neighborhood to estimate optimal fire hydrant placement.

First, consider a local government wants to decide where to build fire hydrants so as to minimize funds spent but maximize the utility of their community. That is, they wish to cover the neighborhood with hydrants so that everyone is close enough so that an average fire trucks hose can extinguish the fire and save the house while not building more hydrants than they need.

This graph was generated in Matlab and though appears to self intersect, we can untwist the graph so that it does not as Matlab only knows how to display it up to isomorphism. Here we have 34 vertices and 63 edges with edges labeled by their respective weightings which here is the actual distance in meters. For covering purposes we find that an average fire house with normal length and pressure can work with 46 meters of distance. The graph connections are geographically based, that is the maximal amount of non overlapping connections were constructed for the graph such that edges did not go through trees, houses, etc. In particular, it was constructed to represent how a fire hose could reach houses in a reasonable manner.

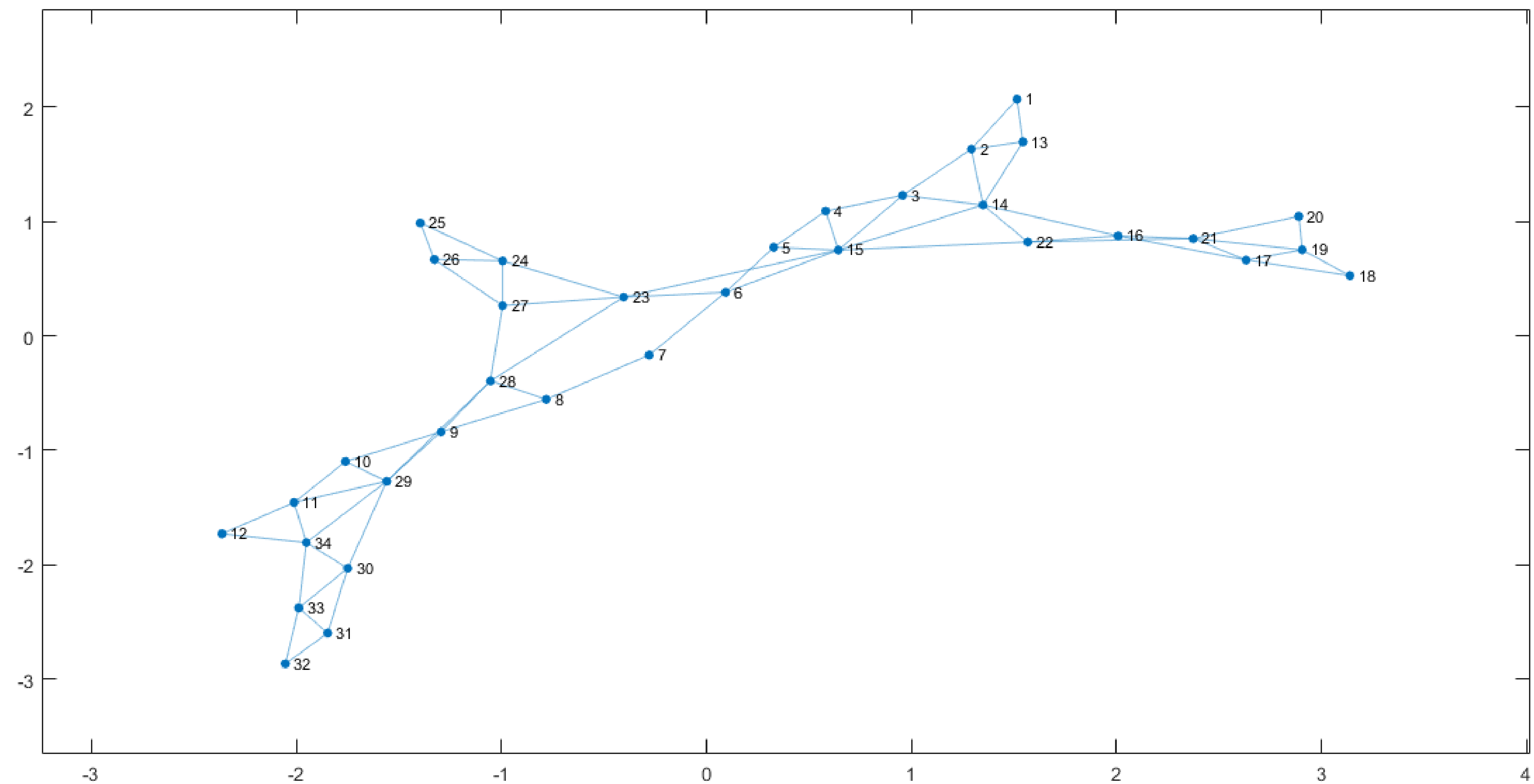

Figure 1: Our neighborhood constructed up to isomorphism, omitting edge weights.

Next we consider the utility structure of the graph. We set everyone to have homogeneous benefit of 153, representing the meters from a fire hydrant can reasonably expect to be able to assist in extinguishing a fire based on city fire codes and let the geodesic be constructed with p=1 and the edge weights determining our penalty. Therefore, those who are very close to a hydrant reap the highest benefit but as long as you are within range you will have a positive utility and those out of range will have a utility of zero.

$$u(i) = 153 - g(j)$$

where j is the node for our selected central placement

Now we construct the adjacency matrix of the graph and solve for the most central vertex. Due to the large size of the adjacency matrix, we will present only an abbreviated version of the complete findings for matrix sections. Here the largest eigenvalue of our graph is 405.8271 which has an eigenvector which ranks vertex 15 as the highest at .5360, thus this node is a good starting point to consider fire hydrant placement. Now we must if this indeed creates a covering of our graph using a shortest path algorithm to throw diameter flags. We allow Matlab to iterate through the nodes starting with node one and find the shortest path to node 1 from node 15 is through the following vertex path: $15 \rightarrow 3 \rightarrow 2 \rightarrow 13 \rightarrow 1$ which is 237 meters in total which is greater than our allowable 153 meters.

Thus we must partition the graph using our Laplacian matrix and Fiedler Partition. Constructing our Laplacian as our degree matrix minus our weighted adjacency matrix, we solve its eigenvalue problem finding the second smallest eigenvalue to be 5.0530 with our two sub-graphs constructed as follows:

Nodes in sub-graph 1: 7,8,9,10,11,12,27,28,29,30,31,32,33,34

Nodes in sub-graph 2: 1,2,3,4,5,6,13,14,15,16,17,18,19,20,21,22,23,24,25,26

Now we resolve for our new central node using the adjacency matrix of each induced sub-graph. Our first bisection sub-graph has its largest eigenvalue at 394.2309 and largest element in the eigenvector is node 9 in our sub-graph with value .6055. Next using the shortest path algorithm we check for diameter flags based on the edge weights. Starting at node one we iterate through and find the shortest path is $9 \rightarrow 3 \rightarrow 2 \rightarrow 1$ which has edge weights totaling 332 which is again greater than 153, thus we must cut it again. For its sister sub-graph we find its largest eigenvalue is 401.3392 with its associated eigenvector having the largest element at node 9 as .5876. We now iterate through shortest paths to check for a covering and starting at node 1 in the induced sub-graph as $9 \rightarrow 3 \rightarrow 2 \rightarrow 7 \rightarrow 1$ which gives us a total of 206 which is greater than 153.To save space it is clear we must iterate this many times and thus the remaining graphs and calculations before the final step will be omitted.

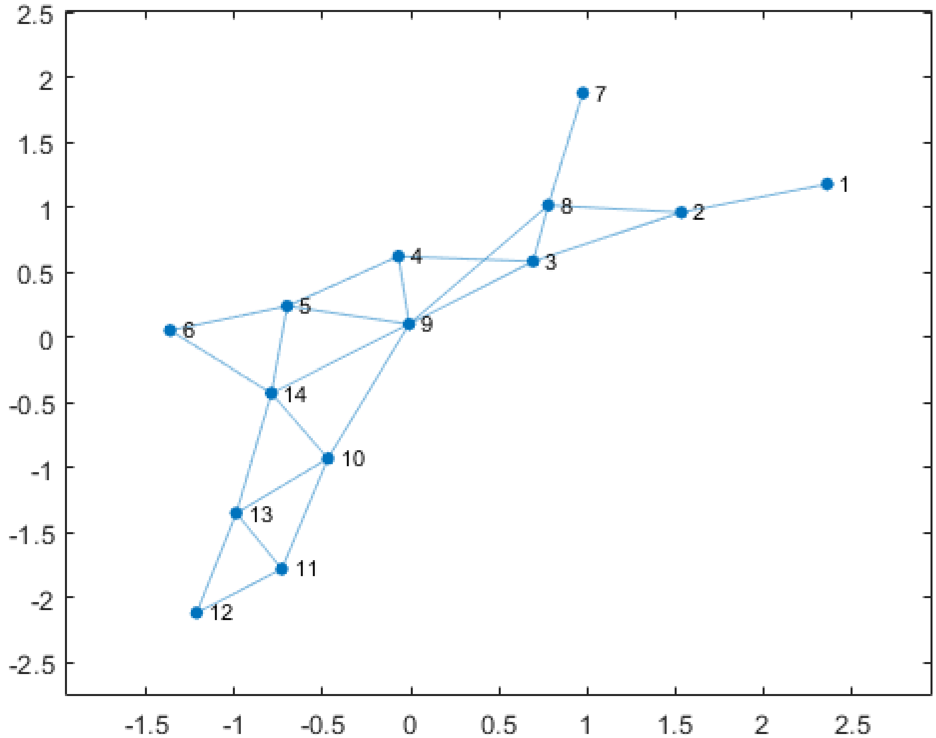

Figure 2: Our first bisection's sub-graph created by the Fiedler Partition

After 13 iterations and naive, non inter step checking we arrive at the answer of 13 fire hydrants being necessary. The placement is as on the following nodes: 34,30,31,8,9,29,11,23,2,15,14,16, and 19. A dendrogram representing this process is included below to exhibit the partitions with a single letter and its double representing sister sub-graphs after a partition.

Next we check this step for a true covering by deleting one selected node from above and seeing if the graph is covered by the rest and then if it is that node is left out, and then the process iterates through the remaining solutions. Doing this we find that choosing nodes 8,9,30, and 31 was redundant and that only 9 nodes were in fact needed. Now we compare this to where the actual fire hydrants are located which is approximately at nodes 2,14,16,5,23,27,29,11,31, and 34, which is 10 nodes. Thus our program differs only at two nodes 5 and 27, which we attribute either to plumbing, zoning issues, as well as this also produces an alternative covering of the area leaving open room for differences. Thus it is shown empirically that our algorithm works and is accurate.

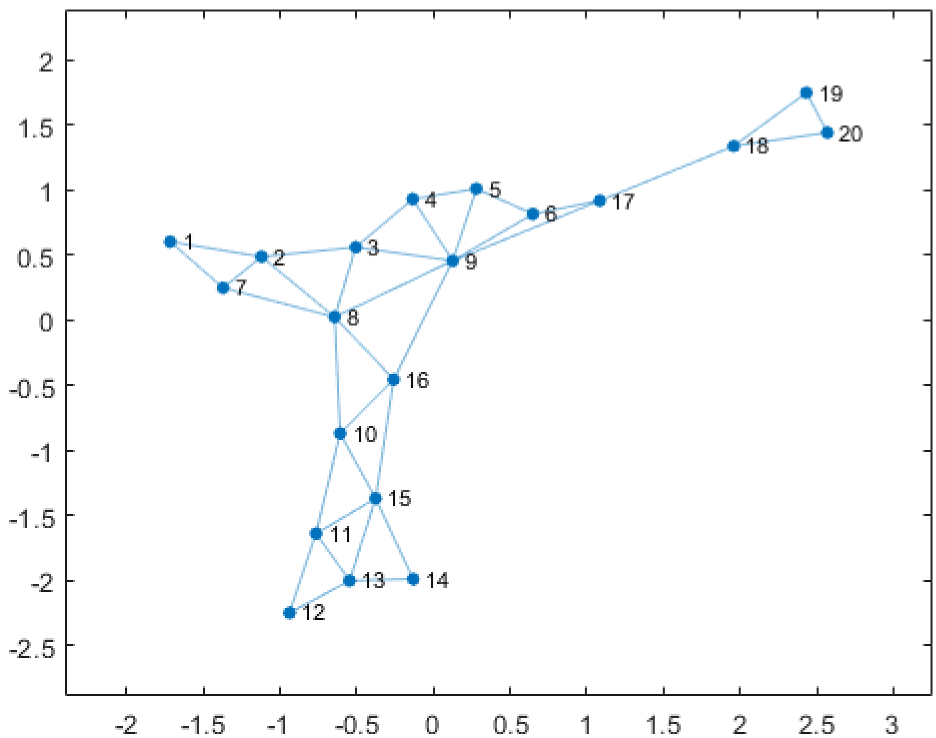

Figure 3: Our first bisection's sister sub-graph created by the Fiedler Partition

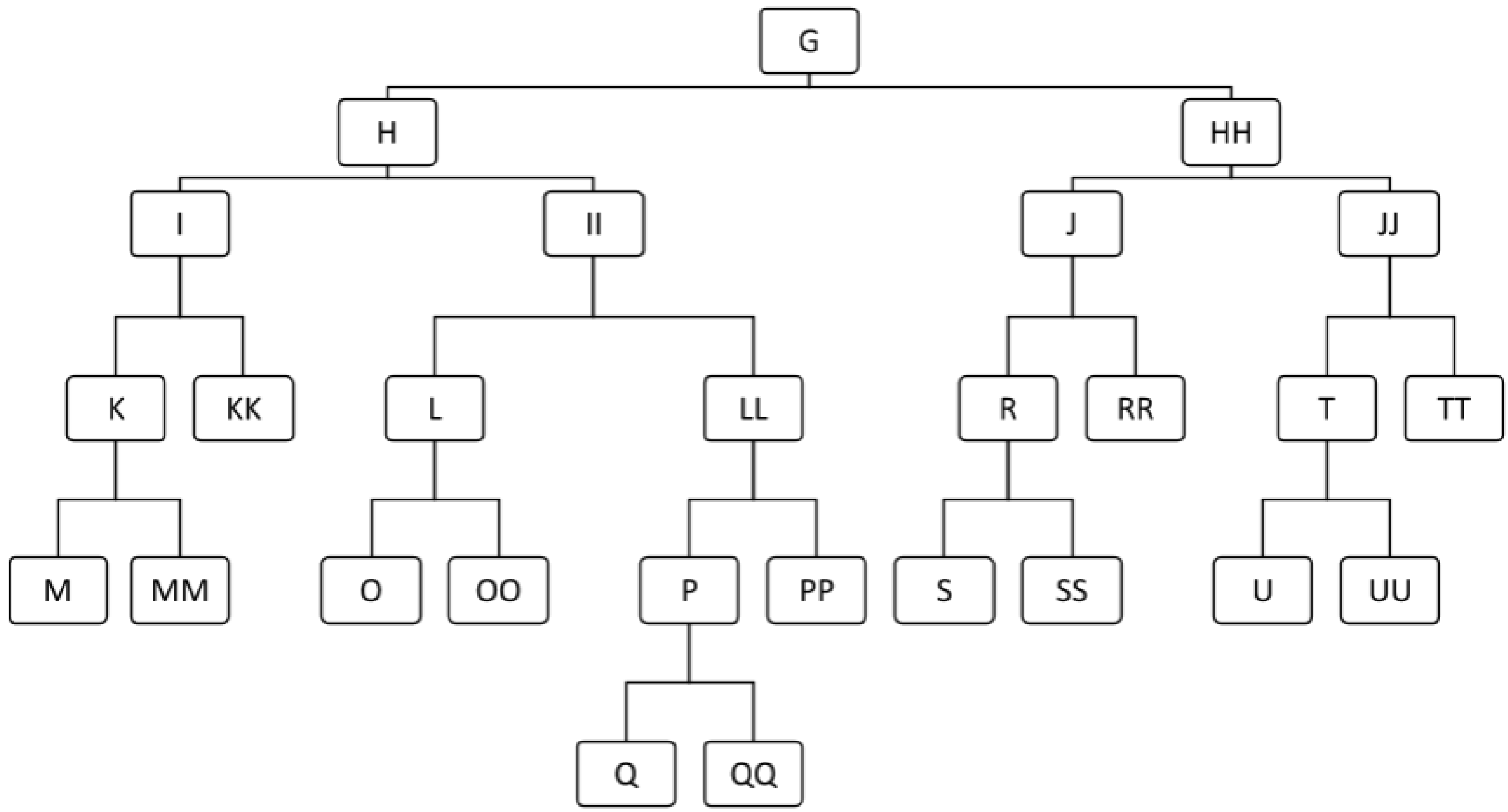

Figure 4: Our partition's dendrogram created by all partitions.

Conclusion

This paper has put forth a method to efficiently allocate public goods along the spatial dimension, mainly with a graph theoretic parameter, which is then verified in placing fire hydrants in a neighborhood. Its strength lies in its range of applicability and relatively simple approach, but has weakness it its time complexity and restriction to planar graphs. Code optimization to decrease its number of numeric calculations will be necessary, and it may be more efficient to approximate the eigenvalue and vectors rather than calculate them explicitly as many direct linear solvers are of an in-optimal time complexity especially when considering large, sparse matrices, such as the ones usually dealt with here.

Further research to be considered can also include the topological properties of graphs and their simplices, that is the generalization of triangles in graphs. Research could include estimating externalities by 'filling in' these triangles based on vertex 'externality emissions' and using some distance metric to perform some process of persistent homology upon them. Other areas could be treating certain goods as a dis-utility and combinatorially optimizing the allocation with respect to this such as with a public park inducing excess noise on nearby homes.

One final thought is the application of this algorithm to dynamic graphs where we allow movement in our vertices or some sort of trade or transition mechanism. This is a much more difficult problem to solve as one must also consider the convergence of the graph and then optimal cutting and centrality.

Refernces

Aingworth, Donald, et al. "Fast estimation of diameter and shortest paths (without matrix multiplication)." SIAM Journal on Computing 28.4 (1999): 1167-1181.

Athreya, Siva, and Rohini Somanathan. "Quantifying spatial misallocation in centrally provided public goods." Economics Letters 98.2 (2008): 201-206.

Barnard, Stephen T., and Horst D. Simon. "Fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems." Concurrency: Practice and experience 6.2 (1994): 101-117.

Bramoullé, Yann, and Rachel Kranton. "Public goods in networks." Journal of Economic Theory 135.1 (2007): 478-494.

Cha, Seung-Jun, et al. "Design and evaluation of experiment methods for improving performance in GIS web services." International Journal of Multimedia and Ubiquitous Engineering 3.1 (2008): 27-43.

Chen, Pin-Yu, and Alfred O. Hero. "Local Fiedler vector centrality for detection of deep and overlapping communities in networks." 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014.

Davern, Michael. "Social networks and economic sociology: a proposed research agenda for a more complete social science." American Journal of Economics and Sociology 56.3 (1997): 287-302.

Dijkstra, Edsger W. "A note on two problems in connexion with graphs." Numerische mathematik 1.1 (1959): 269-271.

Fiedler, Miroslav. "Algebraic connectivity of graphs." Czechoslovak mathematical journal 23.2 (1973): 298-305.

Freeman, Linton C. "A set of measures of centrality based on betweenness." Sociometry (1977): 35-41.

Freeman, Linton C. "Centrality in social networks conceptual clarification." Social networks 1.3 (1978): 215-239.

Grady, Leo, and Eric L. Schwartz. "Isoperimetric graph partitioning for image segmentation." IEEE transactions on pattern analysis and machine intelligence 28.3 (2006): 469-475.

Groves, Theodore, and John Ledyard. "Optimal allocation of public goods: A solution to the free rider problem." Econometrica 45.4 (1977): 783-809.

Henderson, J. Vernon. "The Tiebout model: Bring back the entrepreneurs." Journal of Political Economy 93.2 (1985): 248-264.

Hsieh, Chang-Tai, and Enrico Moretti. Housing constraints and spatial misallocation. No. w21154. National Bureau of Economic Research, 2015.

Jackson, Matthew O., and Alison Watts. "On the formation of interaction networks in social coordination games." Games and Economic Behavior 41.2 (2002): 265-291.

Jackson, Matthew O., and Asher Wolinsky. "A strategic model of social and economic networks." Journal of economic theory 71.1 (1996): 44-74.

Jonker, Roy, and Anton Volgenant. "A shortest augmenting path algorithm for dense and sparse linear assignment problems." Computing 38.4 (1987): 325-340.

Kranton, Rachel E., and Deborah F. Minehart. "A theory of buyer-seller networks." American economic review 91.3 (2001): 485-508.

Montgomery, James D. "Weak ties, employment, and inequality: An equilibrium analysis." American Journal of Sociology 99.5 (1994): 1212-1236.

Newman, Mark EJ. "The structure of scientific collaboration networks." Proceedings of the national academy of sciences 98.2 (2001): 404-409.

Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." Physical review E 69.2 (2004): 026113.

Newman, Mark EJ. "Finding community structure in networks using the eigenvectors of matrices." Physical review E 74.3 (2006): 036104.

Qiu, Huaijun, and Edwin R. Hancock. "Commute times for graph spectral clustering." International Conference on Computer Analysis of Images and Patterns. Springer, Berlin, Heidelberg, 2005.

Raghavan, Usha Nandini, Réka Albert, and Soundar Kumara. "Near linear time algorithm to detect community structures in large-scale networks." Physical review E 76.3 (2007): 036106.

Seidel, Raimund. "On the all-pairs-shortest-path problem." Proceedings of the twenty-fourth annual ACM symposium on Theory of computing. ACM, 1992.

Sieg, Holger, et al. "Estimating the general equilibrium benefits of large changes in spatially delineated public goods." International Economic Review 45.4 (2004): 1047-1077.

Slininger, Brian. "Fiedlers Theory of Spectral Graph Partitioning." (2013).

Szolnoki, Attila, and Matjaž Perc. "Competition of tolerant strategies in the spatial public goods game." New Journal of Physics 18.8 (2016): 083021.

Talen, Emily, and Luc Anselin. "Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds." Environment and planning A 30.4 (1998): 595-613.

Tichy, Noel M., Michael L. Tushman, and Charles Fombrun. "Social network analysis for organizations." Academy of management review 4.4 (1979): 507-519.

Van Driessche, Rafael, and Dirk Roose. "An improved spectral bisection algorithm and its application to dynamic load balancing." Parallel computing 21.1 (1995): 29-48.

Wang, Qiang, Nanrong He, and Xiaojie Chen. "Replicator dynamics for public goods game with resource allocation in large populations." Applied Mathematics and Computation 328 (2018): 162-170.

Watts, Duncan J. "Networks, dynamics, and the small-world phenomenon." American Journal of sociology 105.2 (1999): 493-527.

Wu, Fang, and Bernardo A. Huberman. "Finding communities in linear time: a physics approach." The European Physical Journal B 38.2 (2004): 331-338.